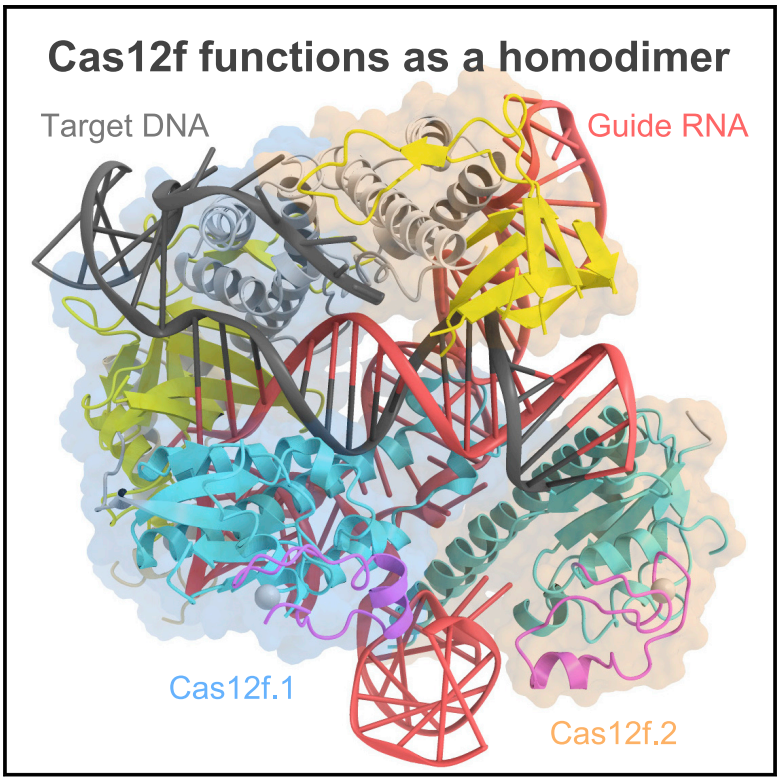
Molecular Cell

Article

Structure of the miniature type V-F CRISPR-Cas effector enzyme

Graphical Abstract



Highlights

- Cryo-EM structure of Cas12f in complex with a guide RNA and its target DNA
- Type V-F effector complex consisting of a Cas12f dimer and a single guide RNA
- Molecular mechanism of the RNA-guided DNA cleavage by the miniature Cas12f
- Evolutionary insights into a possible origin of the type V CRISPR-Cas enzymes

Authors

Satoru N. Takeda, Ryoya Nakagawa, Sae Okazaki, ..., Keitaro Yamashita, Hiroshi Nishimasu, Osamu Nureki

Correspondence

nisimasu@g.ecc.u-tokyo.ac.jp (H.N.), nureki@bs.s.u-tokyo.ac.jp (O.N.)

In Brief

Takeda et al. report the cryo-EM structure of Cas12f in complex with a guide RNA and its target DNA. The structure reveals that two Cas12f molecules bind the guide RNA to form an effector complex and provides mechanistic insights into the RNA-guided DNA cleavage by the miniature Cas12f enzyme.

Takeda et al., 2021, Molecular Cell *81*, 1–13
February 4, 2021 © 2020 Elsevier Inc.
https://doi.org/10.1016/j.molcel.2020.11.035

CellPress

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

Molecular Cell

CellPress

Article

Structure of the miniature type V-F CRISPR-Cas effector enzyme

Satoru N. Takeda,[1] Ryoya Nakagawa,[1] Sae Okazaki,[1,2] Hisato Hirano,[1] Kan Kobayashi,[1,3] Tsukasa Kusakizako,[1] Tomohiro Nishizawa,[1] Keitaro Yamashita,[1,4] Hiroshi Nishimasu,[1,2,*] and Osamu Nureki[1,5,*]

[1]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[2]Present address: Structural Biology Division, Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
[3]Present address: Peptidream Inc., 3-25-23 Tonomachi, Kawasaki-ku, Kawasaki City, Kanagawa Prefecture, 210-0821, Japan
[4]Present address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, UK
[5]Lead contact
*Correspondence: nisimasu@g.ecc.u-tokyo.ac.jp (H.N.), nureki@bs.s.u-tokyo.ac.jp (O.N.)
https://doi.org/10.1016/j.molcel.2020.11.035

SUMMARY

RNA-guided DNA endonucleases derived from CRISPR-Cas adaptive immune systems are widely used as powerful genome-engineering tools. Among the diverse CRISPR-Cas nucleases, the type V-F Cas12f (also known as Cas14) proteins are exceptionally compact and associate with a guide RNA to cleave single- and double-stranded DNA targets. Here, we report the cryo-electron microscopy structure of Cas12f1 (also known as Cas14a) in complex with a guide RNA and its target DNA. Unexpectedly, the structure revealed that two Cas12f1 molecules assemble with the single guide RNA to recognize the double-stranded DNA target. Each Cas12f1 protomer adopts a different conformation and plays distinct roles in nucleic acid recognition and DNA cleavage, thereby explaining how the miniature Cas12f1 enzyme achieves RNA-guided DNA cleavage as an "asymmetric homodimer." Our findings augment the mechanistic understanding of diverse CRISPR-Cas nucleases and provide a framework for the development of compact genome-engineering tools critical for therapeutic genome editing.

INTRODUCTION

CRISPR-Cas systems in bacteria and archaea provide adaptive immunity against foreign nucleic acids and are divided into two classes (classes 1 and 2) and six types (types I–VI) (Hille et al., 2018; Makarova et al., 2020). Class 2 systems include types II, V, and VI and encompass single multidomain effector Cas proteins, such as Cas9 (type II) and Cas12 (type V). Cas9 associates with dual RNA guides (CRISPR RNA [crRNA] and *trans*-activating crRNA [tracrRNA]) or a single guide RNA (sgRNA) and cleaves double-stranded DNA (dsDNA) targets at a sequence complementary to a 20 nt guide segment in the RNA guide and flanked by an NGG (where N is any nucleotide) protospacer adjacent motif (PAM) (Gasiunas et al., 2012; Jinek et al., 2012). Among the diverse type V Cas12 enzymes (Makarova et al., 2020), type V-A Cas12a (also known as Cpf1) binds a crRNA and cleaves dsDNA targets with TTTV (where V is A, G, or C) PAMs (Zetsche et al., 2015). Cas9 contains two nuclease domains, HNH and RuvC, which cleave the target strand (TS) and non-TS (NTS) in dsDNA targets, respectively (Gasiunas et al., 2012; Jinek et al., 2012). In contrast, Cas12a cleaves both the TS and NTS with a single RuvC nuclease domain (Swarts et al., 2017; Swarts and Jinek, 2019). As Cas9 and Cas12a exhibit robust nuclease activities in eukaryotic cells, they are widely used as versatile genome-engineering tools (Cong et al., 2013; Zetsche et al., 2015).

Recent studies identified type V-F Cas12f (formerly Cas14) proteins as exceptionally compact RNA-guided DNA endonucleases (Harrington et al., 2018). Cas12f enzymes consist of ~400–700 amino acid residues and are much smaller than Cas9 and Cas12 (~950–1,400 amino acids). The Cas12f1 (also known as Cas14a1) from an uncultured archaeon consists of 529 residues and lacks detectable sequence identity with any other known proteins, except for the presence of the RuvC domain (Figure S1). Despite its small size, Cas12f1 associates with dual crRNA:tracrRNA guides (which can be combined into an sgRNA) (Harrington et al., 2018) and cleaves dsDNA targets with TTTR (where R is A or G) PAMs (Karvelis et al., 2020). The guide RNA of Cas12f1 lacks sequence similarity with those of the other Cas12 enzymes, such as Cas12a (Zetsche et al., 2015; Yamano et al., 2016), Cas12b (Shmakov et al., 2015; Yang et al., 2016), and Cas12e (also known as CasX) (Burstein et al., 2017; Liu et al., 2019). Therefore, the action mechanism of the miniature type V-F Cas12f nucleases remains enigmatic.

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

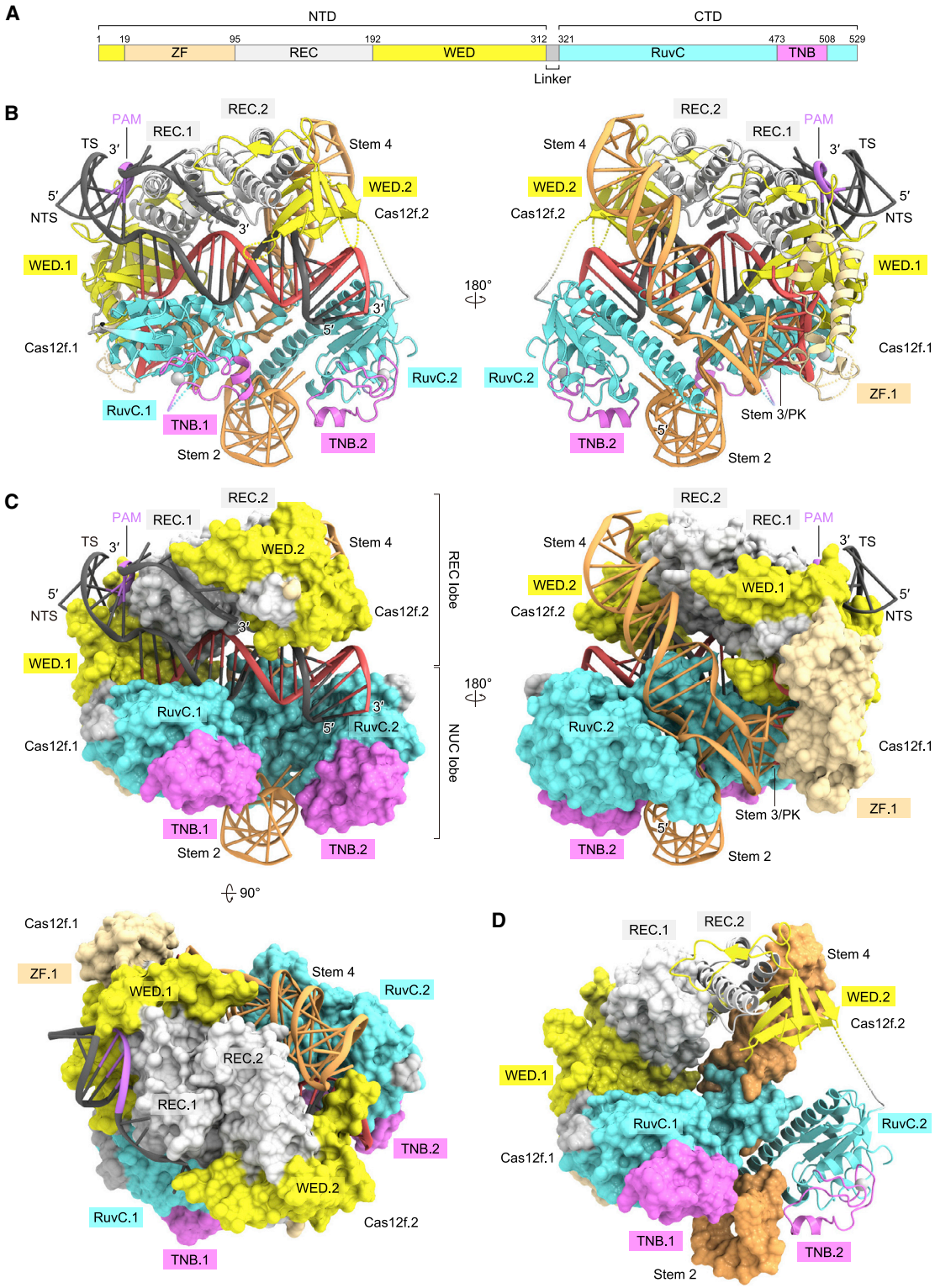


Figure 1. Cryo-EM structure of the Cas12f-guide RNA-target DNA complex

(A) Domain structure of Cas12f. NTD, amino-terminal domain; CTD, carboxy-terminal domain.

(B) Overall structure of the Cas12f-sgRNA-target DNA complex. The zinc ions in the ZF and TNB domains are shown as gray spheres. The disordered regions are indicated as dotted lines. TS, target strand; NTS, non-target strand.

(legend continued on next page)

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

Table 1. Cryo-EM data collection and refinement statistics

Data collection and processing	
Sample	Cas12f-sgRNA-DNA
EMDB ID	EMD-30299
PDB ID	7C7L
Microscope	Titan Krios G3i
Detector	Gatan K3 camera
Magnification	105,000
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	48.7
Defocus range (μm)	−0.8 to −1.6
Pixel size (Å/px)	0.83
Symmetry imposed	*C*1
Number of movies	2,848
Initial particle images	1,960,343
Final particle images	87,253
Map resolution (Å)	3.3
FSC threshold	0.143
Map-sharpening *B* factor (Å^2)	−107.977
Model building and refinement	
Model composition	
Protein atoms	6,828
Nucleic acid atoms	3,317
Metals	3
RMSDs from ideal	
Bond lengths (Å)	0.0074
Bond angles (°)	1.498
Validation	
Clashscore	5.35
Rotamer outliers (%)	8.05
Ramachandran plot	
Favored (%)	90.5
Allowed (%)	9.1
Outlier (%)	0.5

RESULTS

DNA cleavage activity

We examined the *in vitro* DNA cleavage activity of the purified Cas12f1 (hereafter referred to as Cas12f for simplicity) with its sgRNA toward a target dsDNA with a TTTG PAM. Cas12f efficiently cleaved the target DNA at 25–50 mM but not 100–150 mM NaCl concentrations (Figures S2A and S2B), consistent with a previous study (Karvelis et al., 2020). In addition, the D326A mutant did not cleave the target DNA, as previously observed (Harrington et al., 2018; Karvelis et al., 2020), confirming that D326 in the RuvC domain is critical for DNA cleavage (Figures S2A and S2C).

Overall structure

To clarify the Cas12f-mediated DNA cleavage mechanism, we determined the cryoelectron microscopy (cryo-EM) structure of Cas12f (the inactive D326A mutant) in complex with the sgRNA (180 nt) and its target dsDNA (40 bp) with a TTTG PAM, at an overall resolution of 3.3 A (Figures 1A–1D and S3A–S3H; Video S1; Table 1). Surprisingly, the structure revealed that two Cas12f molecules (referred to as Cas12f.1 and Cas12f.2) assemble with one sgRNA molecule to form a ribonucleoprotein effector complex (Figures 1A–1D). Cas12f can be divided into an amino-terminal domain (NTD) and a carboxy-terminal domain (CTD), which are connected by a linker loop. The NTD consists of three domains: the wedge (WED), recognition (REC), and zinc finger (ZF) domains. The CTD consists of the RuvC domain and another ZF domain, termed the target nucleic acid-binding (TNB) domain. The Cas12f dimer adopts a bilobed architecture consisting of a REC lobe and a nuclease (NUC) lobe, with the guide RNA-target DNA heteroduplex bound to the central channel between the two lobes (Figures 1B and 1C). The REC lobe is formed by the WED, ZF, and REC domains of Cas12f.1 (WED.1/ZF.1/REC.1) and Cas12f.2 (WED.2/ZF.2/REC.2), whereas the NUC lobe is formed by the RuvC and TNB domains of Cas12f.1 (RuvC.1/TNB.1) and Cas12f.2 (RuvC.2/TNB.2).

Domain structures

The WED domain comprises a seven-stranded β-barrel flanked by an α helix and a β hairpin and adopts an oligonucleotide/oligosaccharide-binding (OB) fold similar to those of other Cas12 enzymes (Yamano et al., 2016; Yang et al., 2016; Stella et al., 2017; Liu et al., 2019; Swarts and Jinek, 2019), despite their limited sequence similarity (Figure S4A). The ZF and REC domains are inserted between the strands β1 and β2 in the WED domain. The ZF domain contains a CCCH-type ZF, in which a zinc ion is coordinated by C50, H53, C69, and C72 (Figures S4A and S4B). The REC domain comprises four α helices and is much smaller than those of other Cas12 enzymes and thus contributes mainly to the miniaturization of Cas12f (Figure S4A). The RuvC domain has an RNase H fold, consisting of a five-stranded mixed β sheet flanked by four α helices, and D326, E422, and D510 form a catalytic center similar to those of other Cas12 enzymes (Yamano et al., 2016; Yang et al., 2016; Stella et al., 2017; Liu et al., 2019; Swarts and Jinek, 2019) (Figure S4A). The TNB domain is inserted between the strand β5 and the helix α6 in the RuvC domain and contains a CCCC-type ZF, in which a zinc ion is coordinated by C475, C478, C500, and C503 (Figures S4A and S4C). The four cysteine residues are conserved among the Cas12f enzymes (Figure S1). An X-ray fluorescence elemental analysis of the purified Cas12f protein indicated that Cas12f binds zinc ions (Figure S4D).

The TNB domains of the Cas12 enzymes (also known as the Nuc domain for Cas12a and Cas12b and the target-strand loading [TSL] domain for Cas12e) adopt distinct structures (Figure S4A). The TNB domains of Cas12f and Cas12e contain two CXXC ZF motifs, while the TNB domain of Cas12f is smaller

(C and D) Molecular surface of the Cas12f dimer. In (C), the two Cas12f protomers (Cas12f.1 and Cas12f.2) are shown as surface models. In (D), Cas12f.1 and Cas12f.2 are shown as surface and ribbon models, respectively. The guide RNA scaffold is shown as a surface model, whereas the guide segment and target DNA are omitted for clarity.
See also Figures S1–S5.

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

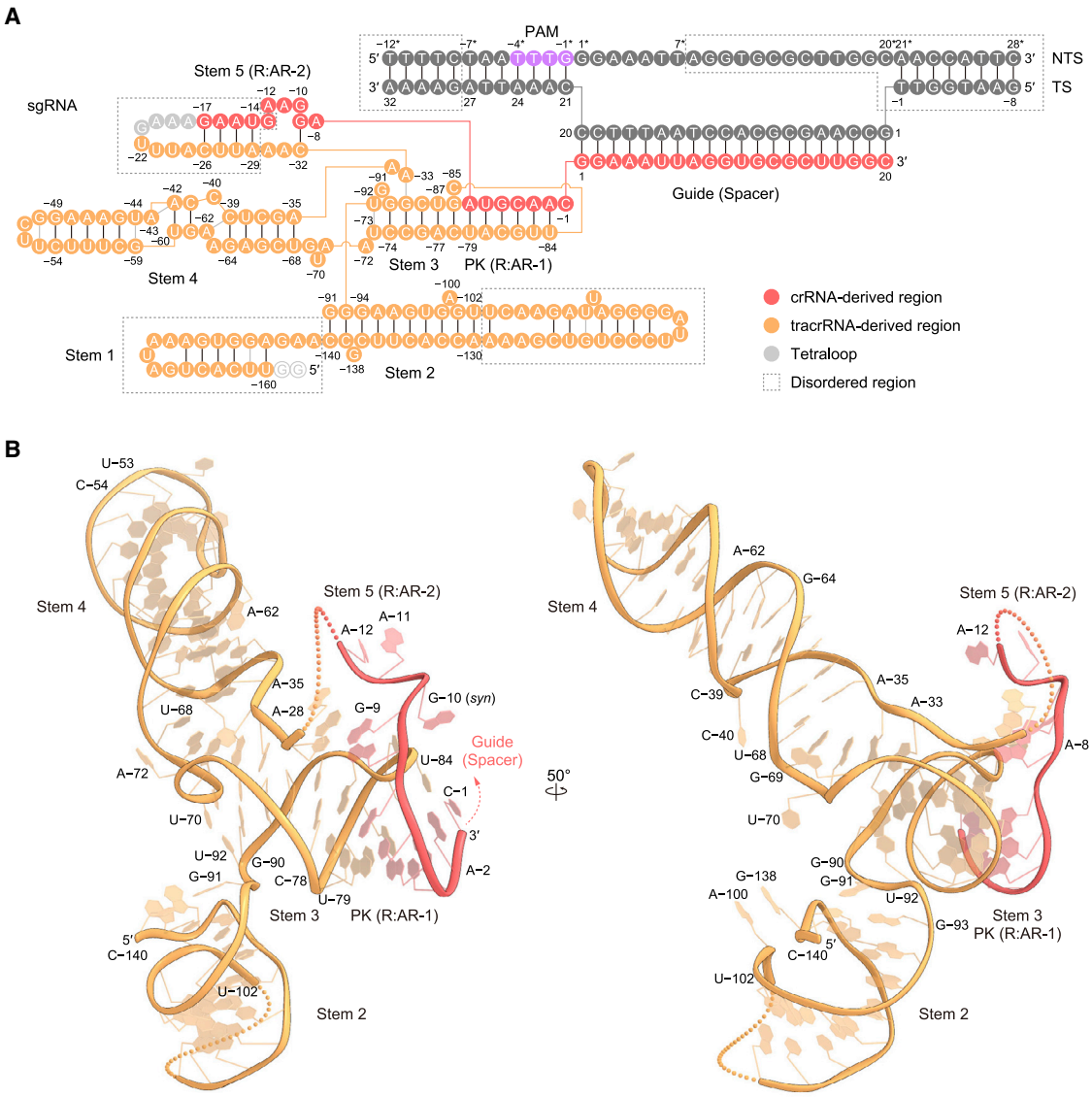


Figure 2. Guide RNA architecture

(A) Schematic of the sgRNA and target DNA. The disordered regions are enclosed in dashed boxes. TS, target strand; NTS, non-target strand; PK, pseudoknot; R:AR, crRNA repeat-tracrRNA anti-repeat duplex.

(B) Structure of the guide RNA scaffold.

See also Figure S6.

than that of Cas12e. The TNB domains of Cas12a and Cas12b have unrelated structures and facilitate target DNA cleavage by the RuvC domain (Yang et al., 2016; Swarts et al., 2017; Swarts and Jinek, 2019). As described previously (Yang et al., 2016; Swarts et al., 2017; Liu et al., 2019; Swarts and Jinek, 2019), these domains adjacent to the RuvC domains probably participate in placement of both the TS and NTS into the RuvC active site. We thus propose that these domains should be collectively referred to as the TNB domain.

A structural comparison of Cas12f.1 and Cas12f.2 revealed a notable difference in the arrangement between their NTD and CTD, which is facilitated by the flexible linker loop and local structural variations in their individual domains (Figures S5A–S5F). ZF.1 (residues 18–93), WED.1 (residues 256–286), and RuvC.1 (residues 368–382) in Cas12f.1 interact with the guide RNA scaffold, whereas the equivalent regions in Cas12f.2 are exposed to the solvent and disordered in the complex structure (Figures 1B–1D and S5A–S5F), indicating that Cas12f undergoes structural changes upon guide RNA binding.

Guide RNA architecture

The sgRNA (U[−160]–C20) consists of the 20 nt guide segment and the 160 nt RNA scaffold, comprising five stems (stems 1–5) and a pseudoknot (PK) (Figures 2A, 2B, S6A, and S6B).

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

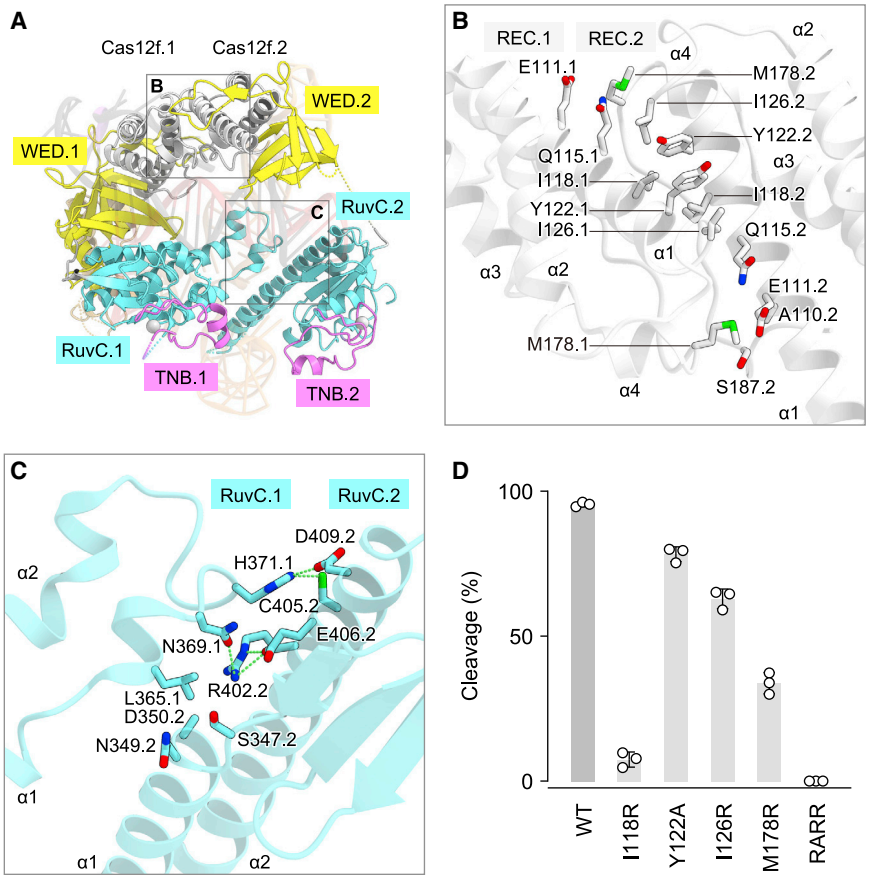


Figure 3. Dimer interfaces

(A) Dimer interfaces between Cas12f.1 and Cas12f.2.

(B) Primary interface between REC.1 and REC.2.

(C) Secondary interface between RuvC.1 and RuvC.2.

(D) *In vitro* DNA cleavage activities of the wild-type (WT) Cas12f and the dimer interface mutants. RARR, I118R/Y122A/I126R/M178R. Data are mean ± SD (n = 3).

See also Figure S7.

Stem 1 (U[−160]–A[−141]) and the upper stem regions of stem 2 (A[−129]–U[−103]) and stem 5 (A[−29]–G[−13]) are disordered in the structure, suggesting the flexibility of these regions. The structure revealed unexpected features of the guide RNA scaffold. First, U(−84)–U(−79) base pair with A(−7)–A(−2) to form the PK (crRNA repeat-tracrRNA anti-repeat duplex 1 [R:AR-1]). The PK coaxially stacks with stem 3 to form a continuous helix. Second, G(−13)–A(−11) do not base pair with the previously predicted C(−26)–U(−28) (Harrington et al., 2018), with A(−12), A(−11), and G(−10) flipped out from the stem, indicating that A(−29)–C(−26), rather than A(−25)–U(−22), base pair with U(−14)–G(−17) to complete stem 5 (R:AR-2). Indeed, the deletion of A(−25)–U(−22) did not affect Cas12f-mediated DNA cleavage (Figure S6C). Third, the sgRNA contains three base triples, G(−89)–C(−75)·A(−33) (stem 3), G(−64)–C(−39)·A(−62) (stem 4a), and U(−60)–A(−42)·A(−43) (stem 4b), which stabilize the RNA scaffold structure (Figure S6D).

Dimer interface

Cas12f dimerizes through two interfaces in an asymmetric manner (Figure 3A). The primary interface is symmetric and formed by hydrophobic residues, I118, Y122, I126, and M178, in REC.1 and REC.2 (Figure 3B). The secondary interface is asymmetric and formed by the α1-α2 loop of RuvC.1 and the helices α1 and α2 of RuvC.2 (Figure 3C). H371.1 and N369.1 hydrogen bond with C405.2/D409.2 and R402.2, respectively, while L365.1 interacts with S347.2, N349.2, and D350.2. The I118R, Y122A, I126R, and M178R mutations reduced the DNA cleavage activities (Figures 3D, S7A, and S7B). Furthermore, the I118R/Y122A/I126R/M178R (RARR) mutant lacked DNA cleavage activity (Figure 3D). In the presence of the sgRNA, the wild-type (WT) Cas12f and the RARR mutant eluted similarly from the size-exclusion column at a position corresponding to a molecular mass of 198 kDa, consistent with the molecular mass of the $(Cas12f)_2$-sgRNA complex (184 kDa), rather than that of the Cas12f-sgRNA complex (121 kDa) (Figure S7C). These data indicated that both the WT Cas12f and the RARR mutant form a dimer bound to the sgRNA, at least under the tested conditions. Together, our structural and functional findings suggested that the correct positioning of Cas12f.2 relative to Cas12f.1 in the $(Cas12f)_2$-sgRNA complex (i.e., the integrity of the dimer interface) is required for the DNA cleavage (in particular, the R-loop formation), but not for the effector complex formation.

Furthermore, we analyzed the oligomeric states of the WT Cas12f and the RARR mutant using size-exclusion chromatography. As a control, we created a dimeric mutant (Dimer), in which two Cas12f molecules are connected by a linker (Figures S7A and S7B). In the absence of the sgRNA, the WT and the RARR mutant eluted from the column later than the Dimer mutant (Figure S7D), suggesting that the guide RNA is required for the Cas12f dimerization.

Guide RNA recognition

The sgRNA is extensively recognized by both Cas12f.1 and Cas12f.2, with the helices α1 and α2 of the RuvC.1 and RuvC.2 playing central roles in the RNA scaffold recognition (Figures 4 and 5A–5F).

Stem 2 is recognized by RuvC.2 mainly through interactions between its lower stem region and the helix α1 of RuvC.2 (Figure 5C). The C(−140)–G(−91) base pair in the basal region of stem 2 stacks with F359.2 and A360.2. G(−138) and A(−100) are sandwiched by K330.2 and F352.2 and form hydrogen

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

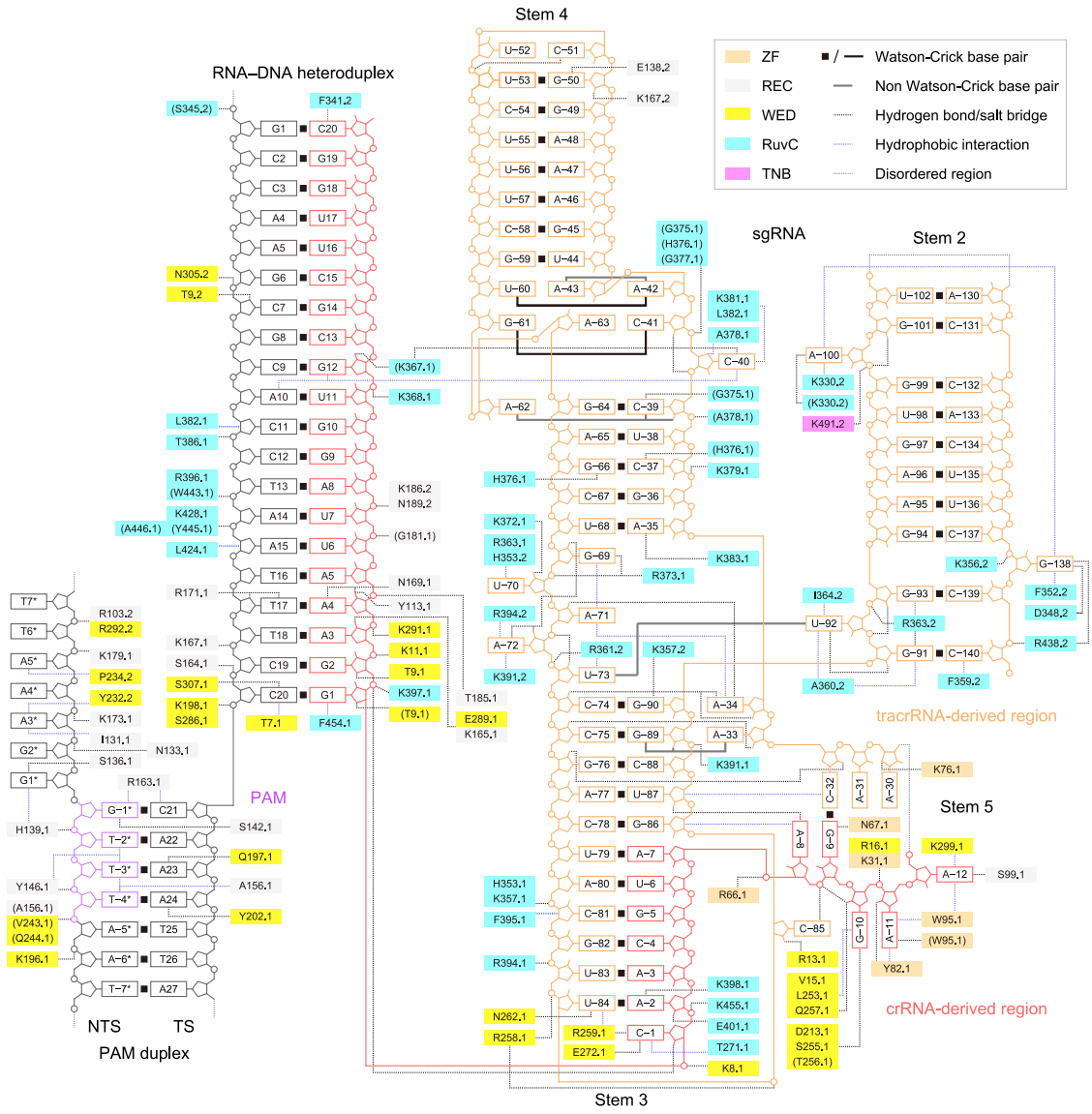


Figure 4. Schematic of nucleic acid recognition

The residues that interact with the nucleic acids through their main chains are shown in parentheses. The disordered regions are indicated as dashed gray lines.

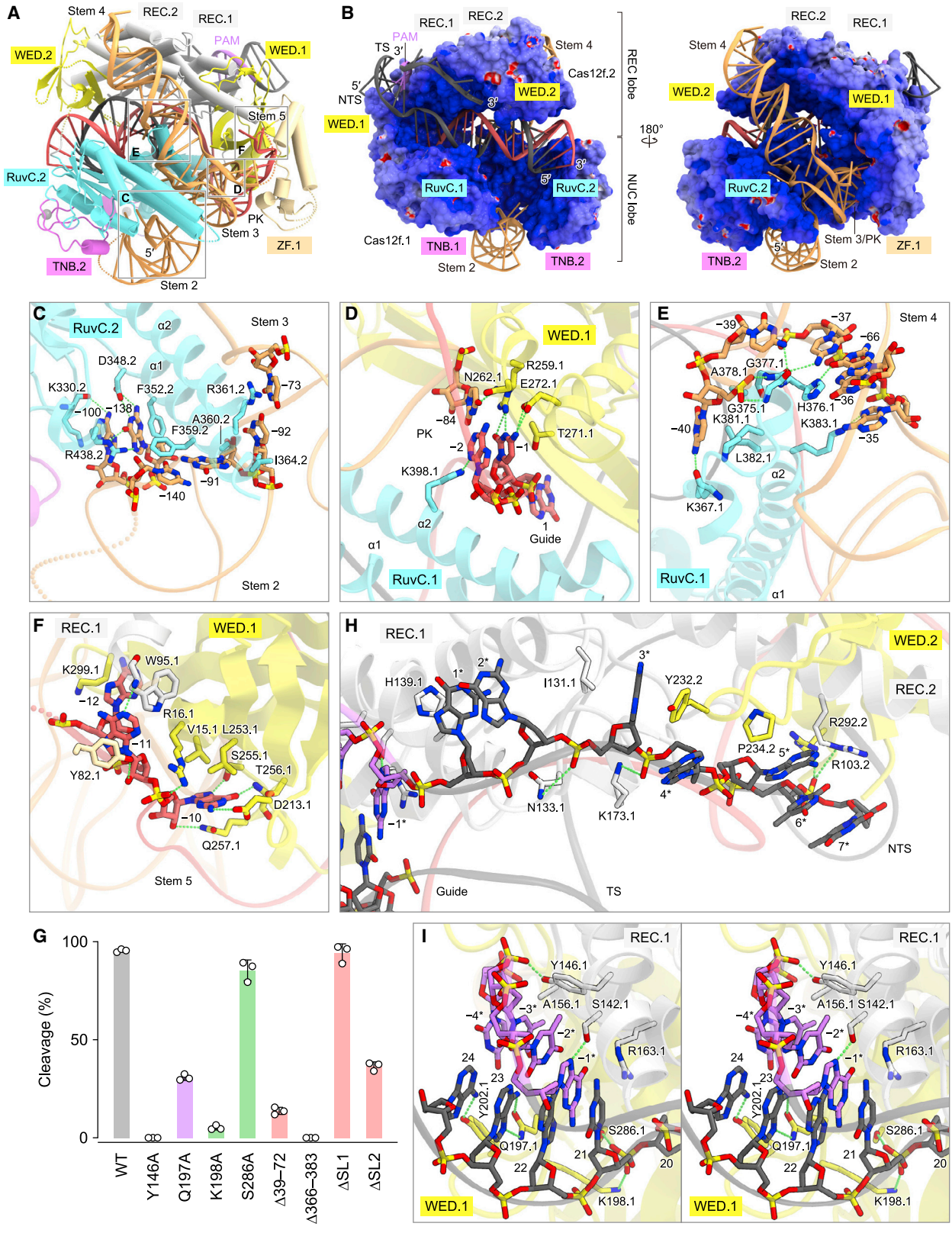
bonds with D348.2/R438.2 and K330.2, respectively. The deletion of stems 1 and 2 (U[−160]–G[−94]), but not stem 1 (U[−160]–A[−144]), reduced the Cas12f-mediated DNA cleavage (Figure 5G), confirming the functional importance of stem 2.

The stem 3-PK helix is recognized by WED.1, ZF.1, and RuvC.1 mainly through interactions with its sugar-phosphate backbone (Figures 4, 5A, and 5B). In addition, the first U(−84)–A(−2) base pair in the PK is recognized by N262.1 and K398.1 through hydrogen bonds, with U(−84) and A(−2) stacking with R259.1 and C(−1), respectively (Figure 5D). C(−1) between the PK and the guide segment is strictly recognized by R259.1, T271.1, and E272.1. The first U(−92)·U(−73) base pair in stem 3 stacks with A360.2, R361.2, and I364.2 of RuvC.2 (Figure 5C).

Stem 4 interacts with RuvC.1 and REC.2 and thus bridges the REC and NUC lobes (Figures 4, 5A, and 5B). The lower stem region of stem 4 (stem 4a) is recognized by the α1-α2 loop in RuvC.1 (Figure 5E). The nucleobases of C(−39), G(−66)–C(−37), and A(−35) in stem 4a form hydrogen bonds with G375.1, H376.1, and K383.1 in the α1-α2 loop, respectively. Notably, C(−40) is flipped out from stem 4 to extensively interact with the side chains of A378.1, K381.1, and L382.1 and the main chains of K367.1, G375.1, and G377.1 (Figure 5E) and thus participates in the RNA-DNA heteroduplex recognition. The deletion of the α1-α2 loop (residues 366–383) abolished the DNA cleavage activity (Figure 5G), while the equivalent region in Cas12f.2 is exposed to the solvent and disordered in the complex structure (Figure S7A). These results indicated that the interaction between RuvC.1 and stem 4 is critical for the Cas12f-mediated

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

CellPress



(legend on next page)

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

DNA cleavage. The upper stem region of stem 4 (stems 4b and 4c) interacts with REC.2 through charge and shape complementarity (Figure 5B).

Stem 5 is recognized by WED.1, ZF.1, and REC.1 (Figures 4, 5A, and 5B). A(−12), A(−11), and G(−10) are flipped out from the stem and sandwiched by W95.1/K299.1, Y82.1/W95.1, and V15.1/L253.1/Q257.1, respectively, with A(−11) forming a hydrogen bond with W95.1 (Figure 5F). In addition, G(−10) adopts the *syn* conformation and forms multiple hydrogen bonds with D213.1, S255.1, and T256.1. The deletion of the ZF motif (residues 39–72) in the ZF domain impaired the DNA cleavage activity (Figure 5G), while ZF.2 is disordered in the structure (Figure S7A), indicating the functional importance of the interaction between ZF.1 and the guide RNA.

Target DNA recognition

The guide RNA-target DNA heteroduplex is accommodated within the positively charged central channel and recognized through interactions with its sugar-phosphate backbone (Figures 4 and 5B), explaining the RNA-dependent DNA recognition mechanism of Cas12f. The seven nucleotides (dG1*–dT7*) in the displaced, single-stranded NTS are recognized by REC.1 and REC.2/WED.2 in a sequence-independent manner. H139.1, I131.1/Y232.2, and P234.2 form stacking interactions with the nucleobases of dG1*, dA3*, and dA5* of the NTS, respectively, while N133.1, K173.1, R103.2, and R292.2 interact with the sugar-phosphate backbone (Figure 5H). We did not observe clear densities for nucleotides (−8)–(−1) and 28–32 of the TS, and nucleotides (−12*)–(−8*) and 8*–28* of the NTS, suggesting their flexibility.

The TTTG PAM-containing duplex is recognized by REC.1 and WED.1 (Figure 5I). The nucleobases of dT(−4*)–dT(−2*) in the PAM form hydrophobic interactions with A156.1 and Y146.1. Y146.1 also interacts with the backbone phosphate group between dT(−4*) and dT(−3*). The nucleobase of dG(−1*) forms hydrogen-bonding and stacking interactions with S142.1 and R163.1, respectively. In addition, the nucleobases of dA24 and dA23, which base pair with dT(−4*) and dT(−3*), form hydrogen bonds with Y202.1 and Q197.1, respectively. The Y146A and Q197A mutations abolished and reduced the DNA cleavage activities, respectively (Figure 5G), while the equivalent residues in Cas12f.2 did not contact the nucleic acids (Figure S7A), confirming the functional importance of Y146.1 and Q197.1 for the PAM recognition. Together, these observations explain the preference of Cas12f for the TTTR PAM.

The backbone phosphate group between dC21 and dC20 in the NTS is recognized by K198.1 and S286.1 in WED.1 (K198.2 and S286.2 are disordered) (Figures 5I and S7A), thereby facilitating the heteroduplex formation, as observed in the other Cas12 enzymes (Yamano et al., 2016; Yang et al., 2016; Stella et al., 2017; Swarts et al., 2017). The K198A and S286A mutants exhibited substantially and slightly reduced cleavage activities, respectively (Figure 5G), suggesting the critical role of K198.1 for the DNA unwinding.

DNA cleavage mechanism

Consistent with a previous study (Karvelis et al., 2020), sequencing of the target DNA cleavage products revealed that Cas12f cleaved the TS and NTS at 24 and 22 nt upstream of the PAM, respectively (Figure 6A). The Cas12 enzymes usually cleave both the TS and NTS at a single RuvC active site, with the TNB (also known as Nuc or TSL) domain facilitating the loading of the TS and NTS into the RuvC active site (Yang et al., 2016; Swarts et al., 2017; Liu et al., 2019; Swarts and Jinek, 2019). In the Cas12f structure, the location of RuvC.1 is similar to those of the RuvC domains of the other Cas12 enzymes, whereas RuvC.2 is close to the 5′ end of the TS (Figure 6B).

To explore the Cas12f-mediated DNA cleavage mechanism, we prepared the D326.1A and D326.2A mutants, in which RuvC.1 and RuvC.2 are selectively inactivated, respectively (Figure 6C). As the Dimer mutant (Figure S7A) can bind to the sgRNA in two different orientations, its two RuvC domains (N-terminal RuvC.1 and C-terminal RuvC.2) can be located at both the RuvC.1 and RuvC.2 positions in the Cas12f-sgRNA-target DNA complex. Thus, the two RuvC active sites in the Dimer mutant cannot be selectively inactivated by the D326A mutations in RuvC.1 and RuvC.2. Notably, residues 366–383 in RuvC.1 participate in the RNA scaffold recognition and are critical for DNA cleavage (Figures 5E and 5G), whereas residues 366–383 in RuvC.2 are exposed to the solvent and disordered in the complex structure (Figure S7A). To selectively inactivate the two RuvC active sites, we connected L365.2 and P384.2 in the Dimer mutant by a linker and deleted residues 366–383 in RuvC.2, to create the DimerΔ mutant (Figure 6C). As expected, the Dimer and DimerΔ mutants exhibited comparable activities, albeit with lower efficiency than that of the WT Cas12f (Figure 6D). These results indicated that the DimerΔ mutant can function only when bound to the sgRNA in a defined orientation, in which the N-terminal RuvC.1 and C-terminal RuvC.2 are located at the RuvC.1 and RuvC.2 positions in the present Cas12f structure, respectively. We then substituted D326.1 and D326.2 in the DimerΔ mutant with alanine to create the D326.1A and D326.2A mutants, respectively (Figure 6C). As the D326.1A and D326.2A mutants can function only when bound to the sgRNA in the defined orientation, RuvC.1 and RuvC.2 are selectively inactivated in the D326.1A and D326.2A mutants, respectively.

Notably, the D326.1A mutant lacked DNA cleavage activity, whereas the D326.2A mutant had activity comparable with that

Figure 5. Recognition of the guide RNA and target DNA

(A) Recognition sites of the guide RNA scaffold.

(B) Electrostatic surface potential of the Cas12f dimer.

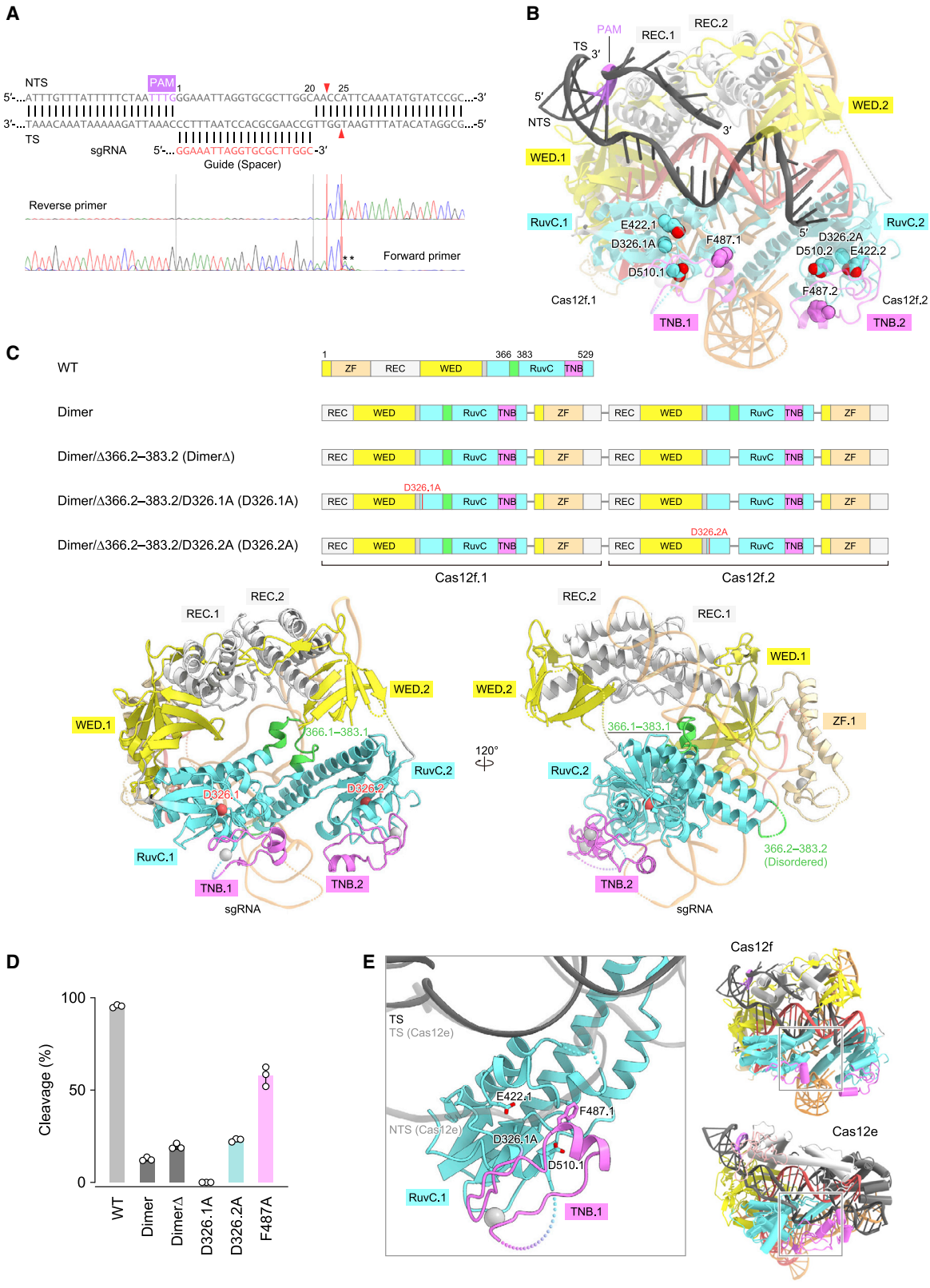
(C–F) Recognition of stems 2/3 (C), PK (D), stem 4 (E), and stem 5 (F).

(G) *In vitro* DNA cleavage activities of the WT Cas12f and the Cas12f mutants with the WT sgRNA and the WT Cas12f with either the stem 1-deleted (ΔSL1) or stems 1 and 2-deleted (ΔSL2) sgRNA. Data are mean ± SD (n = 3).

(H and I) Recognition of the displaced NTS (H) and the PAM duplex (stereo view) (I).

See also Figure S7.

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035



(legend on next page)

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

of the DimerΔ mutant (Figure 6D), suggesting that RuvC.1 cleaves both the TS and NTS. A structural comparison with Cas12e suggested that F487.1 in TNB.1 interacts with the target DNA (Figure 6E). Indeed, the F487A mutant exhibited reduced activity (Figure 6D), suggesting that TNB.1 is involved in the DNA binding, with F487 facilitating the TS recruitment to RuvC.1, as in the other Cas12 enzymes (Yang et al., 2016; Swarts et al., 2017; Liu et al., 2019; Swarts and Jinek, 2019).

DISCUSSION

Unexpectedly, our cryo-EM structure of the Cas12f-sgRNA-target DNA complex revealed that two Cas12f molecules (Cas12f.1 and Cas12f.2) form an asymmetric homodimer and recognize the sgRNA and its target DNA. Our mutational experiments confirmed the functional importance of the observed interactions between Cas12f.1 and Cas12f.2 and between Cas12f.1/Cas12f.2 and the guide RNA. These structural and functional data indicated that the Cas12f dimer bound to a guide RNA is a catalytically competent form of the type V-F CRISPR-Cas effector complex. The two Cas12f protomers exhibit both global and local conformational differences and play unique roles in the nucleic acid recognition and DNA cleavage. Intriguingly, distinct sets of amino acid residues in the WED, REC, and RuvC domains of the two protomers form extensive contacts with different regions of the guide RNA and target DNA. These findings highlight the structural plasticity of ribonucleoprotein complexes. Our structural findings also revealed that Cas12f is the unprecedented class 2 CRISPR-Cas enzyme that, like the class 1 enzymes (Hille et al., 2018; Makarova et al., 2020), forms a multi-protein, rather than a single-protein, effector complex.

Structural comparisons of Cas12f with Cas12a (Swarts and Jinek, 2019), Cas12b (Yang et al., 2016), and Cas12e (Liu et al., 2019) highlighted the mechanistic conservation among the type V Cas12 enzymes, despite their limited sequence similarity (Figure 7). These Cas12 enzymes commonly have a bilobed architecture consisting of the REC and NUC lobes, with the guide RNA-target DNA heteroduplex accommodated within the central channel. Our biochemical analysis suggested that, like the other Cas12 enzymes, Cas12f cleaves both strands of the target DNA at the single RuvC active site, with the adjacent TNB domain helping load the target DNA into the RuvC active site. In some of the other Cas12 structures, the displaced NTS is ordered and bound to a positively charged groove in their RuvC domains (Figure 7). The Cas12a structure contains a nearly complete R-loop, in which NTS nucleotides 1*–13* are ordered, with NTS nucleotides 7*–13* bound in the RuvC domain (Swarts and Jinek, 2019). In the Cas12b structure, a 4 nt segment of the NTS is ordered and bound to the RuvC catalytic pocket, although the intervening segment between nucleotides 1* and 2* and the 4 nt segment is disordered (Yang et al., 2016). The Cas12e structure contains a full R-loop, in which the displaced NTS is bound to the surface groove formed by the RuvC and TNB domains (Liu et al., 2019). In addition, the Cas12b and Cas12e structures in other states revealed that the extended, single-stranded segment of the TS beyond the RNA-DNA heteroduplex folds back to access the RuvC catalytic pocket (Yang et al., 2016). In contrast, the single-stranded segments of the TS and NTS are not bound to the RuvC active site in the present Cas12f structure (Figure 7). These structural differences suggest that the RuvC domain of Cas12f binds the DNA strand less effectively, compared with those of the other Cas12 enzymes.

The structural comparison also revealed the mechanistic diversity among the Cas12 enzymes (Figure 7). These Cas12 enzymes specifically assemble with their cognate guide RNA scaffolds with diverse architectures and recognize T-rich PAMs in distinct manners. Cas12a uses only a crRNA, whereas Cas12b, Cas12e, and Cas12f associate with their cognate RNA scaffolds formed by crRNA and tracrRNA. In Cas12a, the PI domain is inserted into the WED domain, and the TTTV PAM-containing duplex is bound to a channel formed by the REC, WED, and PI domains. In contrast, Cas12b, Cas12e, and Cas12f recognize the TTN, TTCN, and TTTR PAMs in distinct manners, using their REC and WED domains, respectively. The REC lobe of Cas12f is formed by the REC.1 and REC.2/WED.2 domains from the two protomers, whereas those of the other Cas12 enzymes consist of their REC1 and REC2 domains. Notably, the REC lobe (REC.1/REC.2/WED.2) and the RuvC.2 domain in the Cas12f dimer are located at positions similar to the REC1 and REC2 domains in the other Cas12 structures, respectively. These structural findings suggest that the larger monomeric Cas12 enzymes gained additional domains over time, and evolved from the compact dimeric Cas12f enzymes.

Whereas the type V-F Cas12f nuclease failed to protect against invading plasmids in bacterial cells, probably because of its salt sensitivity, some type V-U3 Cas12f proteins efficiently interfered with the plasmid transformation (Karvelis et al., 2020), suggesting that they could be harnessed for eukaryotic genome editing. Therefore, the present structure of Cas12f, together with future structural studies and protein engineering of the other miniature type V-F and V-U3 enzymes, will pave the way for the development of compact genome-engineering tools.

Figure 6. DNA cleavage mechanism

(A) Cleavage sites in the target DNA. The plasmid target with the TTTG PAM was cleaved by the Cas12f-sgRNA complex at 50°C for 10 min, and the cleavage products were then analyzed by Sanger sequencing. The cleavage sites are marked by red triangles. The additional non-templated adenines are indicated by asterisks. TS, target strand; NTS, non-target strand.

(B) Active sites of Cas12f.1 and Cas12f.2.

(C) Design of the D326.1A and D326.2A mutants. The N- and C-terminal Cas12f genes in the Dimer mutant were codon-optimized for *E. coli* and human, respectively, for mutagenesis.

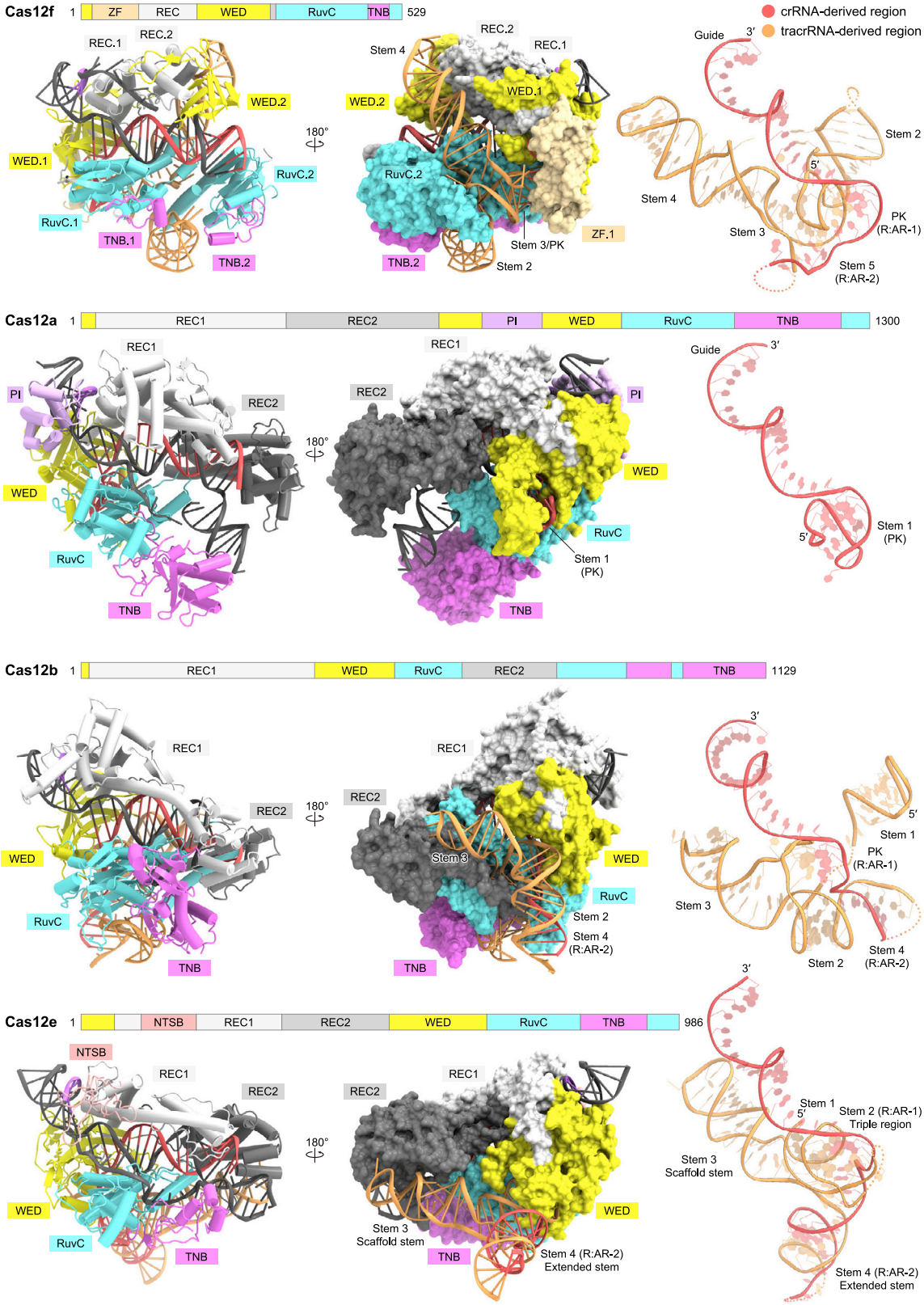
(D) *In vitro* DNA cleavage activities of the WT Cas12f and the RuvC mutants. Data are mean ± SD (n = 3).

(E) Active site of Cas12f.1. The TS and NTS in the Cas12e structure (PDB: 6NY2) are shown in semi-transparent gray and superimposed on the Cas12f structure. The Cas12e protein and the guide RNA are omitted for clarity.

See also Figure S7.

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

CellPress



(legend on next page)

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

Limitations of study

The Cas12f-guide RNA-target DNA complex structure revealed that Cas12f forms a dimer that binds to the guide RNA and its target DNA, thereby explaining the RNA-guided DNA cleavage mechanism of the miniature Cas12f enzyme. A structural comparison of Cas12f with the other Cas12 enzymes highlighted the mechanistic conservation and divergence between them, thus providing insights into a possible origin of the type V Cas12 enzymes. Our data indicated that Cas12f requires the guide RNA for dimerization, and that the Cas12f dimer bound to the guide RNA is a catalytically competent effector complex. Nonetheless, the assembly mechanism of the effector complex remains unclear. In addition, the DNA recognition and cleavage mechanisms of Cas12f are not fully understood, as the target DNA strands are not visible in the active site formed by the RuvC and TNB domains in the present structure. Thus, structural studies of Cas12f in different states, such as the Cas12f-guide RNA complex and the catalytically active Cas12f R-loop complex, and more compact type V-U3 Cas12f enzymes will be required to clarify the action mechanisms of the miniature Cas12f family enzymes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
  - Experimental model and subject details
- METHOD DETAILS
  - Protein and RNA preparation
  - Electron microscopy sample preparation
  - Electron microscopy data collection and processing
  - Model building and validation
  - X-ray fluorescence analysis
  - Size-exclusion chromatography
  - *In vitro* cleavage experiment
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2020.11.035.

ACKNOWLEDGMENTS

We thank Mr. Masahiro Hiraizumi for assistance with the cryo-EM analysis and Drs. Y. Kawano and K. Hasegawa for assistance with the X-ray fluorescence analysis. The synchrotron radiation experiments were performed at the BL41XU of SPring-8 with the approval of JASRI proposal 2019B0153. This research was supported by the Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research [BINDS]) from the Agency for Medical Research and Development (AMED) under grant JP20am0101115. This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI grants 18H02384 (H.N.) and 18K19284 (H.N.), Japan Science and Technology Agency (JST) CREST grant JPMJCR19H5 (H.N.), and AMED grant JP19am0401005 (H.N. and O.N.).

AUTHOR CONTRIBUTIONS

S.N.T. performed biochemical experiments with assistance from R.N. and cryo-EM analysis with assistance from K.K., T.K., and T.N. S.O. prepared samples. S.N.T., H.H., K.Y., and H.N. performed model building and structural refinement. H.N. conceived the project. S.N.T. and H.N. wrote the manuscript with help from all authors. H.N. and O.N. supervised the research.

DECLARATION OF INTERESTS

O.N. is a co-founder, board member, and scientific advisor for Modalis and Curreio. S.N.T., H.N., and O.N. have filed a patent application related to this work.

Received: July 8, 2020
Revised: October 9, 2020
Accepted: November 19, 2020
Published: December 16, 2020

REFERENCES

Afonine, P.V., Klaholz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., Adams, P.D., and Urzhumtsev, A. (2018a). New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallogr. D Struct. Biol. *74*, 814–840.

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018b). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544.

Burstein, D., Harrington, L.B., Strutt, S.C., Probst, A.J., Anantharaman, K., Thomas, B.C., Doudna, J.A., and Banfield, J.F. (2017). New CRISPR-Cas systems from uncultivated microbes. Nature *542*, 237–241.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U S A *109*, E2579–E2586.

Harrington, L.B., Burstein, D., Chen, J.S., Paez-Espino, D., Ma, E., Witte, I.P., Cofsky, J.C., Kyrpides, N.C., Banfield, J.F., and Doudna, J.A. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. Science *362*, 839–842.

Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. Cell *172*, 1239–1259.

Figure 7. Comparisons of the diverse type V CRISPR-Cas12 enzymes

Structural comparisons of Cas12f with Cas12a (Swarts and Jinek, 2019) (Cas12a from *Francisella novicida*) (PDB: 6I1K), Cas12b (Yang et al., 2016) (Cas12b from *Alicyclobacillus acidoterrestris*, also known as C2c1) (PDB: 5U33), and Cas12e (Liu et al., 2019) (Cas12e from *Deltaproteobacteria*, also known as CasX) (PDB: 6NY2).

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Karvelis, T., Bigelyte, G., Young, J.K., Hou, Z., Zedaveinyte, R., Budre, K., Paulraj, S., Djukanovic, V., Gasior, S., Silanskas, A., et al. (2020). PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage. Nucleic Acids Res. *48*, 5016–5023.

Liu, J.J., Orlova, N., Oakes, B.L., Ma, E., Spinner, H.B., Baney, K.L.M., Chuck, J., Tan, D., Knott, G.J., Harrington, L.B., et al. (2019). CasX enzymes comprise a distinct family of RNA-guided genome. Nature *566*, 218–223.

Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. *18*, 67–83.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

Newville, M. (2013). Larch: an analysis package for XAFS and related spectroscopies. J. Phys. Conf. Ser. *430*, 012007.

Nicholls, R.A., Tykac, M., Kovalevskiy, O., and Murshudov, G.N. (2018). Current approaches for the fitting and refinement of atomic models into cryo-EM maps using CCP-EM. Acta Crystallogr. D Struct. Biol. *74*, 492–505.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. *333*, 721–745.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol. Cell *60*, 385–397.

Stella, S., Alcón, P., and Montoya, G. (2017). Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage. Nature *546*, 559–563.

Swarts, D.C., and Jinek, M. (2019). Mechanistic insights into the *cis*- and *trans*-acting DNase activities of Cas12a. Mol. Cell *73*, 589–600.e4.

Swarts, D.C., van der Oost, J., and Jinek, M. (2017). Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a. Mol. Cell *66*, 221–233.e4.

Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. eLife *5*, e17219.

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I.M., Li, Y., Fedorova, I., Nakane, T., Makarova, K.S., Koonin, E.V., et al. (2016). Crystal structure of Cpf1 in complex with guide RNA and target DNA. Cell *165*, 949–962.

Yang, H., Gao, P., Rajashankar, K.R., and Patel, D.J. (2016). PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. Cell *167*, 1814–1828.e12.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759–771.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ *6*, 5–17.

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Cas12f (Cas12f1, also known as Cas14a1)	Harrington et al., 2018	N/A
Cas12f, various mutants	This paper	N/A
Deposited data		
Cas12f D326A coordinates	This paper	PDB: 7C7L
Cas12f D326A EM map	This paper	EMDB: EMD-30299
Cas12f D326A EM raw images	This paper	EMPIAR-10585
In vitro DNA cleavage experiments	This paper; Mendeley Data	https://doi.org/10.17632/r7fgtnxcb5.1
Experimental models: organisms/strains		
E. coli Mach	Thermo Fisher Scientific	C862003
E. coli Rosetta 2 (DE3)	Novagen	71397
Oligonucleotides		
DNA primers	This paper	Table S1
DNA oligos (for structure determination)	This paper	Table S1
Cas12f sgRNA	Harrington et al., 2018	Table S1
Cas12f sgRNA, various mutants	This paper	Table S1
Recombinant DNA		
pE-SUMO-Cas12f	This paper	https://benchling.com/s/seq-Jp0TFd9iPVLeIvMJ3BS2/edit
pE-SUMO-Cas12f, various mutants	This paper	Dimer mutant; https://benchling.com/s/seq-8ZHLwmtijAyp5Vbrkorh/edit
pUC119-T20	This paper	https://benchling.com/s/seq-FniQ3mcf8gLSFomZmUld/edit
Software and algorithms		
SerialEM	Mastronarde, 2005	https://bio3d.colorado.edu/SerialEM/
MotionCor2	Zheng et al., 2017	https://emcore.ucsf.edu/ucsf-software
Relion	Zivanov et al., 2018	https://www3.mrc-lmb.cam.ac.uk/relion/index.php?title=Main_Page
CTFFIND4	Rohou and Grigorieff, 2015	https://grigoriefflab.umassmed.edu/ctffind4
COOT	Emsley et al., 2010; Nicholls et al., 2018	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Rosetta	Wang et al., 2016	https://www.rosettacommons.org/software
PHENIX	(Afonine et al., 2018b)	https://www.phenix-online.org/
REFMAC	Nicholls et al., 2018	https://www.ccp4.ac.uk/
MolProbity	Chen et al., 2010	https://www.phenix-online.org/documentation/reference/molprobity_tool.html
UCSF-Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/
CueMol	N/A	http://www.cuemol.org
Larch Python library	Newville, 2013	https://xraypy.github.io/xraylarch/
Prism 7	GraphPad	https://www.graphpad.com/scientific-software/prism/
Other		
Amicon Ultra-4 Centrifugal Filter Units - 10,000 NMWL	Millipore	UFC801024
Ni-NTA Superflow	QIAGEN	30450

(*Continued on next page*)

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HiTrap Heparin HP	GE Healthcare	17040701
Superdex 200 Increase 10/300 GL	GE Healthcare	28990944
Gel Filtration Calibration Kit	GE Healthcare	28403841
300 mesh R 1/1 holey carbon copper	Quantifoil	https://www.quantifoil.com/products/quantifoil/quantifoil-circular-holes/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Osamu Nureki (nureki@bs.s.u-tokyo.ac.jp).

Materials availability

All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

The structural model, density map, and raw images have been deposited in the Protein Data Bank, the Electron Microscopy Data Bank, and the Electron Microscopy Public Image Archive under the accession codes 7C7L, EMD-30299, and EMPIAR-10585, respectively. The data of unprocessed image files have been deposited in the Mendeley Data repository (https://doi.org/10.17632/r7fgtnxcb5.1).

Experimental model and subject details

Escherichia coli Mach (Thermo Fisher Scientific) and *E. coli* Rosetta2 (DE3) (Novagen) were cultured at 37°C in LB medium (containing 20 mg/l kanamycin) for plasmid preparation and protein expression, respectively.

METHOD DETAILS

Protein and RNA preparation

The gene encoding Cas12f (Cas12f1 from an uncultured archaeon, also known as Cas14a; residues 1–529) was synthesized (Invitrogen GeneArt) and cloned into the modified pE-SUMO vector (LifeSensors), which lacks the SUMO-coding region. The mutations were introduced by a PCR-based method, and the sequences were confirmed by DNA sequencing. The N-terminally His_6-tagged Cas12f proteins were expressed in *E. coli* Rosetta2 (DE3). The *E. coli* cells were cultured at 37°C until the OD_{600} reached 0.8, and the protein expression was then induced by the addition of 0.1 mM isopropyl β-D-thiogalactopyranoside (Nacalai Tesque). The *E. coli* cells were further cultured at 20°C overnight, and harvested by centrifugation. The *E. coli* cells were resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 20 mM imidazole, 1 M NaCl, and 1 mM DTT), lysed by sonication, and then centrifuged. The supernatant was mixed with 1 mL Ni-NTA Superflow resin (QIAGEN), and the mixture was loaded into a Poly-Prep Column (Bio-Rad). The protein was eluted with buffer B (20 mM Tris-HCl, pH 8.0, 0.3 M imidazole, 0.3 M NaCl, and 1 mM DTT). The protein was then loaded onto a 1 mL HiTrap Heparin HP column (GE Healthcare), equilibrated with buffer C (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, and 1 mM DTT). The protein was eluted with a linear gradient of 0.3–2 M NaCl. The purified proteins were stored at −80°C until use. The sgRNAs were transcribed *in vitro* with T7 RNA polymerase, and purified by 10% denaturing (7 M urea) polyacrylamide gel electrophoresis.

Electron microscopy sample preparation

The Cas12f-sgRNA-target DNA complex was reconstituted by mixing the purified Cas12f D326A mutant, the 180-nucleotide sgRNA (180 nucleotides plus 5′ GG for *in vitro* transcription), the 40-nucleotide target DNA strand (Sigma-Aldrich), and the 40-nucleotide non-target DNA strand (Sigma-Aldrich), at a molar ratio of 1:1.2:1.5:1.5. The Cas12f-sgRNA-target DNA complex was purified by size-exclusion chromatography on a Superdex 200 Increase 10/300 column (GE Healthcare), equilibrated with buffer D (20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT). The purified complex solution (~3 mg/ml, 5.4 μl) was mixed with 0.6 μL of $ZnCl_2$ (10 μM final concentration), and the samples (3 μl) were then applied to freshly glow-discharged Cu/Rh 300 mesh R1/1 grids (Quantifoil) on both sides, in a Vitrobot Mark IV (FEI) at 4°C with a waiting time of 10 s and a blotting time of 4 s under 100% humidity conditions. The grids were plunge-frozen in liquid ethane cooled at liquid nitrogen temperature.

Electron microscopy data collection and processing

The cryo-EM data were collected using a Titan Krios G3i microscope (Thermo Fisher Scientific), running at 300 kV and equipped with a Gatan Quantum-LS Energy Filter (GIF) and a Gatan K3 Summit direct electron detector in the electron counting mode. Each movie

Please cite this article in press as: Takeda et al., Structure of the miniature type V-F CRISPR-Cas effector enzyme, Molecular Cell (2020), https://doi.org/10.1016/j.molcel.2020.11.035

was recorded at a nominal magnification of 105,000 × , corresponding to a calibrated pixel size of 0.83 A (The University of Tokyo, Japan) at the electron exposure of 15.8 e^{-}/pix/sec for 2.6 s, resulting in an accumulated exposure of 48.7 e^{-}/A^{2}. The data were automatically acquired by the image shift method using the SerialEM software (Mastronarde, 2005), with a defocus range of −0.8 to −1.6 μm, and 2,848 movies were acquired. The dose-fractionated movies were subjected to beam-induced motion correction and dose-weighting, using the MotionCor2 (Zheng et al., 2017) algorithm implemented in RELION-3 (Zivanov et al., 2018), and the contrast transfer function (CTF) parameters were estimated using CTFFIND4 (Rohou and Grigorieff, 2015).

The data processing was performed using RELION-3. From the 2,848 motion-corrected and dose-weighted micrographs, 1,960,343 particles were initially picked, and extracted at a pixel size of 3.28 A. These particles were subjected to several rounds of 2D and 3D classifications. The selected 143,063 particles were then re-extracted at a pixel size of 1.05 A and subjected to 3D refinement, per-particle defocus refinement, beam-tilt refinement, Bayesian polishing (Zivanov et al., 2019), and 3D classification. The selected 87,253 particles were subjected to 3D refinement, and subsequent postprocessing of the map improved its global resolution to 3.3 A, according to the Fourier shell correlation (FSC) = 0.143 criterion (Rosenthal and Henderson, 2003). The local resolution was estimated by RELION-3.

Model building and validation

The model was manually built using COOT (Emsley et al., 2010; Nicholls et al., 2018), and the protein model was then rebuilt using Rosetta (Wang et al., 2016) against the density map. The model was refined using phenix.real_space_refine ver. 1.16 (Afonine et al., 2018b) and REFMAC5.8 (Nicholls et al., 2018), with the secondary structure and base pair/stacking restraints. The structure validation was performed using MolProbity (Chen et al., 2010) from the PHENIX package. Residues 1–2, 40–46, 57–60, 440–442, 504–507, and 524–529 of Cas12f.1 and residues 1–4, 18–93, 196–204, 210–213, 256–286, 295–299, 312–317, 368–382, and 523–529 of Cas12f.2, nucleotides (−160)–(−141), (−129)–(−103), and (−29)–(−13) of the sgRNA, nucleotides (−8)–(−1) and 28–32 of the TS, and nucleotides (−12*)–(−8*) and 8*–28* of the NTS are not included in the final model, since these regions are not well resolved in the density map. The curve representing model versus full map was calculated using phenix.mtriage (Afonine et al., 2018a), based on the final model and the full, filtered, and sharpened map. The statistics of the 3D reconstruction and model refinement are summarized in Table 1. The cryo-EM density maps were calculated with UCSF Chimera (Pettersen et al., 2004), and molecular graphics figures were prepared with CueMol (http://www.cuemol.org).

X-ray fluorescence analysis

X-ray fluorescence spectra were collected on BL41XU at SPring-8 (Hyogo, Japan), using an XR-100 FAST SDD (Amptek) and a digital pulse processor PX5 (Amptek). The frozen samples were irradiated with an 18.00-keV X-ray beam, and X-ray fluorescence spectra were accumulated for 80 s and 290 s from the Cas12f protein sample and a blank (buffer C), respectively. X-ray emission line data were obtained using the Larch Python library (Newville, 2013).

Size-exclusion chromatography

To examine the Cas12f-sgRNA complex formation, the purified Cas12f protein (2 μM WT or RARR) and the sgRNA (1 μM) were mixed in buffer D (400 μl) and then analyzed on a Superdex 200 Increase 10/300 column, equilibrated with buffer D. The gel-filtration column was calibrated using Gel Filtration Calibration Kit (GE Healthcare). To examine the dimer formation, the purified Cas12f protein (2 μM WT, RARR, or Dimer) was analyzed on a Superdex 200 Increase 10/300 column, equilibrated with buffer E (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT).

In vitro cleavage experiment

The DNA cleavage activity of Cas12f was measured by *in vitro* DNA cleavage experiments. The Cas12f-sgRNA complex (500 nM) was prepared by mixing the purified Cas12f (1 μM) and the sgRNA (1 μM) at 50°C for 3 min, in 10 μL buffer F (5 mM Tris-HCl, pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT). The pre-assembled Cas12f-sgRNA complex (2 μl, 500 nM, 100 nM final concentration) was mixed with the linearized plasmid target containing the 20-nucleotide target sequence and the TTTG PAM (8 μl, 100 ng, 5 nM final concentration), and incubated at 50°C in 10 μL reaction buffer (5 mM Tris-HCl, pH 7.5, 25–150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT). The reaction was stopped by the addition of quench buffer, containing EDTA (20 mM final concentration) and Proteinase K (40 ng). Aliquots (2 μl) were taken at 0.5, 1, 2, and 5 min, and mixed with the quench buffer (6 μl). The reaction products were then analyzed using a MultiNA microchip electrophoresis system (SHIMADZU). To determine the Cas12f DNA cleavage sites, the linearized plasmid target (5 nM) was incubated at 50°C for 10 min with the Cas12f-sgRNA complex (100 nM) in buffer F (50 μl). The reaction mixture was combined with the quench buffer, and then purified with a Wizard SV Gel and PCR Clean-Up System (Promega). The purified cleavage products were analyzed by DNA sequencing (Eurofins Genomics). *In vitro* cleavage experiments were performed at least three times.

QUANTIFICATION AND STATISTICAL ANALYSIS

In vitro cleavage experiments were performed at least three times. Data are shown as mean ± SD (n = 3). Kinetics data were fitted with a one-phase exponential association curve, using Prism 7 (GraphPad).